The nucleotide sequence of *Xenopus laevis* transferrin mRNA

J.E.Moskaitis, R.L.Pastori and D.R.Schoenberg*

Department of Pharmacology, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Rd Bethesda, MD 20889-4799, USA

Submitted September 14, 1990

EMBL accession no. X54530

The three proteins principally involved in the regulation of iron metabolism are transferrin, the transferrin receptor and ferritin. Transferrin is the principal iron transport protein present in vertebrate plasma. It consists of a single polypeptide chain of approximately 80 kDa with two iron binding domains (1). We have isolated a cDNA encoding the complete sequence of transferrin mRNA from a library prepared from male *Xenopus laevis* liver. The sequence of *Xenopus* transferrin mRNA and its derived peptide sequence are presented below. *Xenopus* transferrin is similar in size to the corresponding molecules from human (1) and chicken (2) and shares the same repeated domain structure. Human and chicken transferrin have 2 and 3 sites of N-linked glycosylation, respectively. The glycosylation sites are located in the C-terminal half of these proteins. There are no sites for N-linked glycosylation in *Xenopus* transferrin. *Xenopus* transferrin has 47% and 46% amino acid identity with the chicken

and human proteins, and 63% similarity to both. This degree of homology is significantly less than that found for *Xenopus* and human ferritin, which share 68% amino acid identity and 84.6% overall similarity (3).

ACKNOWLEDGEMENTS

Supported by grants GM38277 from the NIH and CO7577 from the USUHS.

REFERENCES

1. MacGillivray, R.T.A., Mendez, E., Shewale, J.G., Sinha, S.K., Lineback-Zins, J. and Brew, K. (1983) *J. Biol. Chem.* **258**, 3543–3553.
2. Jeltsch, J.-M. and Chambon, P. (1982) *Eur. J. Biochem.* **122**, 291–295.
3. Moskaitis, J.E., Pastori, R.L. and Schoenberg, D.R. (1990) *Nucl. Acids Res.* **18**, 2184.

```

1 TTTACCTTCCAGCCCTGAATAATCCAAGGCAATGGATTTTCTCCGTGTGGCACTGTGCCTGAGCATGCTGGCCCTGTGCCTGGCAATACAAAAAGAAAGCAGGTGCGCTGGTGT
   M D F S L R V A L C L S N L A L C L A I Q K E K Q V R W C
120 GTAAATCAACAGTGAACCTCAAAAGTGCAAGATCTGGTAGACACGTGCAAAAAAGGAAATCAAATGTCTGCTAGAAAAATCAATACTGATGAGTGTCTACTGCTATTCAGG
   V K S N S E L K K C K D L V D T C K N K E I K L S C V E K S N T D E C S L L F R
240 AAGACCATGCAGATGCAATTTGTGTGGACGGGAGGGATGTGTACAAAGGATCACTGCAACCTTATAACCTGAAGCCAATCATGGCAGAAACTATGGTTACATACAGAAACGGACACA
   K T M Q M Q F V W T G G D V Y K G S L Q P Y N L K P I M A E N Y G S H T E T D T
360 TGTACTATGCAGTGGCAGTTGTAATAAATCCAGCAAAATTTACATTTGATGAGCTGAAAGATAAAAAATCATGCCATACCGGAATTGGCAAGACTGCTGGATGGAATATCATCATTTGGA
   C Y Y A V A V V K K S S K F T F D E L K D K K S C H T G I G K T A G W N I I I G
480 TTACTCCTGGAGAGAAAGCTGTTGAAATGGGACGGGCCAGATTCAGAACTTGGAGAAACCGGGTTTCAAAATTCCTCAAAGCCAGCTGTGTTCCCGGAGCCAAAGAACCTAAATTTGTC
   L L L E R K L L K W A G P D S E T W R N A V S K F F K A S C V P G A K E P K L S
600 CAGCTGTGTCCGGAATTAAGGAGCATAAGTGCTCAGCTCTAACACGAGCCGCTACTACACTACGCTGGTCAATCAAGTGTGTTGCAAGATGACAGGGAGATGTGCTTTTGTGAAG
   Q L C A G I K E H K C S R S N H E P Y Y N Y A G A F K C L Q D D Q G D V A F V K
720 CAATCAACTGTACAGAGGAATTCATAAAGATTATGAGCTGCTGTGCCGGGATAATACAGAAAAAGCATAAAGGAATATAAAATTTGCAACTTGGCTAAGTTCTGCTCATGCGGTG
   Q S T V P E E F H K D Y E L L C P D N T R K S I K E Y K N C N L A K V P A H A V
840 TTGACCAAGGGGAGATGATAAATCTAAGGACATCATTGAATTCCTCAAGAGCGCAGAAACACAAGATGTAATTTGTTGAGGCTCCCGGATGGGGAAGGGATCTAATTTTCAA
   L T R G R D D K S K D I I E F L Q E A Q K A K T Q E C K L F R L P G M G K G S N F Q
960 GGACGCGCAGTGAGTCTTATCCCTCCCATCTTCTATGAGCAGTTTCTGTCTCCTCGGAGCCGATTATTCATGATCAAGCACTGAAGGAAGGGGTGAAGGAGGATGACTCTGG
   G Q R S E S Y S P P I F Y G Q D T F S V P R S R L F Q C I Q A L K E G V K E D D S A
1080 GCTCAAGTGAAGTTCGTTGGTGACACAAAGCAAGGCAAGAAACCAAGTGTGATGACTGGACAATATTAGTGGTGGCCATTGAATGCACTGAGGCCCTCCACTGCTGAAGATGT
   A Q V K V R W C T Q S K A E K T K C D D W T T I S G G A I E C T E A S T A E E C
1200 ATCGTACAGATTCTGAAAGGTGATGCTGACGAGTCACTTTGGATGGAGGCTACATGTACACTGCAGGGCTATGTGGTTGGTTCAGTAATGGGAGGATGACTACGACCAAGATGACCTA
   I V Q I L K G D A D A V T L D G G Y M Y T A G L C G L V P V M G E Y Y D Q D D L
1320 ACACCATGCCACGGAGTGTTCACAGGCAAAAGGTATATATTGCTGTAGCCATTGTAAAGAGGGAAACACAAGTCTCATGGAGCAATCTTAGAGGAGTAAGACATGTACACTGCT
   T P C Q R S C S Q A K G V Y Y A V A I V K K G T Q V S W S N L R G V K T C H T A
1440 GTTGACGTACAGTGGCTGGAACATACCGTGGGACTAATACAGTGAACCGGCAAGTGTGATTTGCTAGTTACGTGGGTGAAAGCTGTGCTCCCGGTTCTGATGTGAATCTAAT
   V G R T A G W N I P V G L I T S E T A N C D F A S Y V G E S C A P G S D V K S N
1560 CTGTGCGCGCTGTGATTGGTGATCCGGAAGCTGTCCGAGAGAGAAAAAGTGTTCCTCCAGTGCAGTGAGGCATATTATGGATACAGCGGTGCATTCAGATGTTGGTCCGAGAAA
   L C A L C I G D P E K L S E R E K K C S P S A S E A Y Y G Y S G A F R C L V E K
1680 GGGCAAGTGGGTTTGCAAAACACACCGGATTTGAGAACAGATGGGAAACCCCTGCAGGCTGGGCAAAAGATCGAAATCAGAAGATTTGAGTTGCTGTGCCCCGATGGCAGC
   G Q V G F A K H T V F E N T D G K N P A G W A K D L K S E D F E L L C P D G S
1800 AGAGCGCCGCTCACTGATTATAAGAGGTGCAACCTTGCAGAGGTTCTGCCCCATGCTGTTGTCACCTTGCAGACAGAGGAGCAAGTTCGCAAAATTTGTTGTCATCAACAGTCATTG
   R A P V T Y K R C N L A E V P A H A V T L P D K R E Q V A K I V N Q Q S L
1920 TATGGCCGAAAGGATTCAAAAAGATATCTCCAGATGTTTCAGTCACTGGTGGCAAGACCTTTGTTTAAAGATCCACCCAGTGCTTACTGGAGATTCCAAGCAAAACCAATG
   Y G R K G F Q K D I F Q M F Q S T T G G K D L L F K D S T Q C L L E I P S K T T M
2040 CAAGAATTCCTTGGAGATAAATACACTGCTGTACCAGCCTGAATAAATGCAGACCTCAATGAGGCCTCCTGGCTGCCTGCACAATTCACAGCTGCATGAAGATTATATCATG
   Q E F L G D K Y H T A V T S L N K C S T S N E A S W L P A Q F H S C M K I Y I M
2160 GTAGATTGCTCCCTTTCATCATATAACCCCAACCTCAGTTATTCTCAGATGTGAAGGGGGAATTTTCTTATCTTCTGTTGTTTGTGTTTCAATAAAGTTACTACAT
   V D C P L S I I
2279 GA(A)n

```

* To whom correspondence should be addressed